

Bayes linear variance adjustment for time series

D. J. Wilkinson*

February 5, 2008

Abstract

This paper exhibits quadratic products of linear combinations of observables which identify the covariance structure underlying the univariate locally linear time series dynamic linear model. The first- and second-order moments for the joint distribution over these observables are given, allowing Bayes linear learning for the underlying covariance structure for the time series model. An example is given which illustrates the methodology and highlights the practical implications of the theory.

Keywords: BAYES LINEAR METHODS; DYNAMIC LINEAR MODELS; EXCHANGEABILITY; IDENTIFIABILITY; VARIANCE ESTIMATION.

1 Introduction

In [10], new methodology is developed for the revision of covariance structures underlying two-step invertible dynamic linear models (DLMs). Two-step invertible DLMs are essentially models without a trend component. Here, the locally linear DLM will be discussed, which does have a trend component, and so is not two-step invertible. Interest will be focussed on univariate DLMs, since it is simpler to explain the theory in the univariate context. However, a full covariance matrix approach may be taken to the multivariate counterpart, thus generalising the work in [10]. It will be shown how one may learn about the three different kinds of variance associated with the locally linear dynamic linear model, using partial prior specification for certain aspects of the model. The theory will be applied to the modelling of sales of a particular product from a wholesale depot, and the importance of good estimation of all of the variance components will be demonstrated in the context of this example.

2 Variance modelling

2.1 The linear model

First consider the model for the time series $\{X_1, X_2, \dots\}$.

$$X_t = M_t + Y_{1t}, \quad \forall t \geq 1 \tag{1}$$

$$M_t = M_{t-1} + N_t + Y_{2t}, \quad \forall t \geq 2 \tag{2}$$

$$N_t = N_{t-1} + Y_{3t}, \quad \forall t \geq 2 \tag{3}$$

where beliefs about M_1 and N_1 are specified *a priori*, and the collection of quantities $\{Y_{jk} | j = 1, 2, 3, k \geq 1\}$ have expectation zero, and are mutually uncorrelated. Further, $\text{Var}(Y_{jk}) = v_j$ does not depend on k for $j = 1, 2, 3, k \geq 1$. Using the terminology of [8], this is a second-order description of the univariate locally linear time series DLM. Here interest focusses on Bayes linear methods for learning about the covariance structure underlying this model. Explicitly, the desire is to make inferences about the variances of Y_{1t} , Y_{2t} and Y_{3t} , in order that we might revise specifications for v_1 , v_2 and v_3 .

*E-mail: d.j.wilkinson@durham.ac.uk, WWW: <http://fourier.dur.ac.uk:8000/djw.html>

2.2 The quadratic model

Form the unobservable vector time series $\{\mathbf{Z}_1, \mathbf{Z}_2, \dots\}$, where $\mathbf{Z}_t = (Y_{1t}^2, Y_{2t}^2, Y_{3t}^2)^T$, $\forall t \geq 1$. This time series is judged to be a second-order exchangeable sequence (that is, mean, variance and covariance specifications for the sequence remain invariant under an arbitrary permutation of the sequence), as described in [3]. Using the second-order exchangeability representation theorem [3], we may decompose the vectors, \mathbf{Z}_t in the following way.

$$\mathbf{Z}_t = \mathbf{V} + \mathbf{S}_t, \quad \forall t \geq 1 \quad (4)$$

where $E(\mathbf{S}_t) = 0$, $Cov(\mathbf{V}, \mathbf{S}_t) = 0$, $\forall t$ and $Cov(\mathbf{S}_s, \mathbf{S}_t) = 0 \forall s \neq t$. The additional assumptions, $Cov(V_i, V_j) = 0$, $\forall i \neq j$ and $Cov(S_{it}, S_{jt}) = 0 \forall i \neq j, \forall t$, where $\mathbf{V} = (V_1, V_2, V_3)^T$ and $\mathbf{S}_t = (S_{1t}, S_{2t}, S_{3t})^T, \forall t$, are made. Note that the components of \mathbf{V} represent the underlying variances for Y_{1t} , Y_{2t} and Y_{3t} , and in particular, that $E(\mathbf{V}) = (\text{Var}(Y_{1t}), \text{Var}(Y_{2t}), \text{Var}(Y_{3t}))^T$. Revised beliefs about \mathbf{V} will lead to revised specifications for v_1 , v_2 and v_3 . Next, observables are constructed which are predictive for \mathbf{V} .

3 State independent observables

3.1 Linear observables

First construct the one-step differenced time series, $\{X'_2, X'_3, \dots\}$, where

$$X'_t = X_t - X_{t-1} \quad (5)$$

$$= N_t + Y_{2t} + Y_{1t} - Y_{1(t-1)}, \quad \forall t \geq 2 \quad (6)$$

Next construct the one-, two-, and three-step differences of the differenced series, $\{X_3^{(1)}, X_4^{(1)}, \dots\}$, $\{X_4^{(2)}, X_5^{(2)}, \dots\}$ and $\{X_5^{(3)}, X_6^{(3)}, \dots\}$, where

$$\begin{aligned} X_t^{(1)} &= X'_t - X'_{t-1} \\ &= Y_{3t} + Y_{2t} - Y_{2(t-1)} + Y_{1t} - 2Y_{1(t-1)} + Y_{1(t-2)}, \quad \forall t \geq 3 \end{aligned} \quad (7)$$

$$\begin{aligned} X_t^{(2)} &= X'_t - X'_{t-2} \\ &= Y_{3t} + Y_{3(t-1)} + Y_{2t} - Y_{2(t-2)} + Y_{1t} - Y_{1(t-1)} - Y_{1(t-2)} + Y_{1(t-3)}, \quad \forall t \geq 4 \end{aligned} \quad (8)$$

$$\begin{aligned} X_t^{(3)} &= X'_t - X'_{t-3} \\ &= Y_{3t} + Y_{3(t-1)} + Y_{3(t-2)} + Y_{2t} - Y_{2(t-3)} + Y_{1t} - Y_{1(t-1)} - Y_{1(t-3)} + Y_{1(t-4)}, \\ &\quad \forall t \geq 5 \end{aligned} \quad (9)$$

Note that these observable series only involve the error structure. Also note that the $\{X_3^{(1)}, X_4^{(1)}, \dots\}$ series is (second-order) 3-step exchangeable, as defined in [10] and discussed more fully in [9]. Briefly, a collection of random quantities is (second-order) n -step exchangeable if the expectation and covariance structure over them is invariant under a reflection or arbitrary translation of the collection, and if the covariance between any two members of the collection is fixed provided only that they are a distance of at least n apart. Note similarly that the $\{X_4^{(2)}, X_5^{(2)}, \dots\}$ series is 4-step exchangeable and that the $\{X_5^{(3)}, X_6^{(3)}, \dots\}$ series is 5-step exchangeable.

3.2 Quadratic observables

Form the series of the squares of the linear series, $\{X_3^{(1)2}, X_4^{(1)2}, \dots\}$, $\{X_4^{(2)2}, X_5^{(2)2}, \dots\}$ and $\{X_5^{(3)2}, X_6^{(3)2}, \dots\}$. Note that due to assumptions of (second-order) exchangeability for the quadratic residuals, these series have the same n -step exchangeability properties as the linear series they are constructed from. Consequently, the (second-order) n -step exchangeability representation theorem [10] tells us that the series *identify* the Cauchy limit of the partial arithmetic means. A random quantity is identified by a collection of observables if as much uncertainty as is desired may be resolved by observing an increasing number of the observables.

Lemma 1 *The following identification results hold.*

- $\{X_t^{(1)2} | \forall t \geq 3\}$ identify $6V_1 + 2V_2 + V_3$

- $\{X_t^{(2)^2} | \forall t \geq 4\}$ identify $4V_1 + 2V_2 + 2V_3$
- $\{X_t^{(3)^2} | \forall t \geq 5\}$ identify $4V_1 + 2V_2 + 3V_3$

Proof

Using (4) and (7), the collection $\{X_t^{(1)^2} | \forall t \geq 3\}$ identifies

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=3}^N X_t^{(1)^2} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=3}^N (Y_{3t} + Y_{2t} + Y_{1t} - 2Y_{1(t-1)} - Y_{2(t-1)} + Y_{1(t-2)})^2 \quad (10)$$

$$= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=3}^N (Y_{3t}^2 + Y_{2t}^2 + Y_{1t}^2 + 4Y_{1(t-1)}^2 + Y_{2(t-1)}^2 + Y_{1(t-2)}^2) \quad (11)$$

$$= V_3 + V_2 + V_1 + 4V_1 + V_2 + V_1 \quad (12)$$

$$= 6V_1 + 2V_2 + V_3 \quad (13)$$

Using (8) and (9), the other results follow similarly. \square

In fact, $\{X_t^{(n)^2} | \forall t \geq n+2\}$ identifies $4V_1 + 2V_2 + nV_3$, $\forall n \geq 2$. Note also that

$$\begin{pmatrix} 6 & 2 & 1 \\ 4 & 2 & 2 \\ 4 & 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & -1 & 1/2 \\ -1 & 7/2 & -2 \\ 0 & -1 & 1 \end{pmatrix} \quad (14)$$

and so Lemma 1 inverts to give

Theorem 1 *The following identification results hold.*

- $\{\frac{1}{2}X_t^{(1)^2} - X_t^{(2)^2} + \frac{1}{2}X_t^{(3)^2} | \forall t \geq 5\}$ identify V_1
- $\{-X_t^{(1)^2} + \frac{7}{2}X_t^{(2)^2} - 2X_t^{(3)^2} | \forall t \geq 5\}$ identify V_2
- $\{-X_t^{(2)^2} + X_t^{(3)^2} | \forall t \geq 5\}$ identify V_3

Proof

Clear from Lemma 1 and (14). \square

Of course, the partial arithmetic means of these collections may be used as frequentist unbiased estimators of the underlying variances (though they would not necessarily have optimum variance properties). Here, the quadratic observables will be used in order to allow Bayes linear updating of beliefs for the underlying variances.

4 Bayes linear methods

A Bayes linear approach is taken to subjective statistical inference, making expectation (rather than probability) primitive. An overview of the methodology is given in [1]. The emphasis of this paper is on learning about underlying means; however, the foundations of the theory are quite general, and are outlined in the context of second-order exchangeability in [4], and discussed for more general situations in [5]. Bayes linear methods may be used in order to learn about any quantities of interest, provided only that a mean and variance specification is made for all relevant quantities, and a specification for the covariance between all pairs of quantities is made. No distributional assumptions are necessary. There are many interpretive and diagnostic features of the Bayes linear methodology. These are discussed with reference to [B/D] (the computer language used for the analysis of the example given in this paper) in [6].

5 Covariance structure over the quadratic observables

In order to carry out Bayes linear updating of the underlying variances, the covariances over and between the underlying variables and the predictive observables are required. They are given as follows:

Lemma 2

$$\text{Cov}(V_1, X_t^{(1)2}) = 6\text{Var}(V_1), \quad \forall t \geq 3 \quad (15)$$

$$\text{Cov}(V_1, X_t^{(2)2}) = 4\text{Var}(V_1), \quad \forall t \geq 4 \quad (16)$$

$$\text{Cov}(V_1, X_t^{(3)2}) = 4\text{Var}(V_1), \quad \forall t \geq 5 \quad (17)$$

$$\text{Cov}(V_2, X_t^{(1)2}) = 2\text{Var}(V_2), \quad \forall t \geq 3 \quad (18)$$

$$\text{Cov}(V_2, X_t^{(2)2}) = 2\text{Var}(V_2), \quad \forall t \geq 4 \quad (19)$$

$$\text{Cov}(V_2, X_t^{(3)2}) = 2\text{Var}(V_2), \quad \forall t \geq 5 \quad (20)$$

$$\text{Cov}(V_3, X_t^{(1)2}) = 1\text{Var}(V_3), \quad \forall t \geq 3 \quad (21)$$

$$\text{Cov}(V_3, X_t^{(2)2}) = 2\text{Var}(V_3), \quad \forall t \geq 4 \quad (22)$$

$$\text{Cov}(V_3, X_t^{(3)2}) = 3\text{Var}(V_3), \quad \forall t \geq 5 \quad (23)$$

Proof

These are a trivial consequence of Lemma 1 and the n -step exchangeability representation theorem. \square

The covariances between the quadratic observables themselves are rather complex, and are given in the appendix.

6 Bayes linear adjustment for the variances

Theorem 1 shows that Bayes linear fitting of the underlying variances, V_1 , V_2 , and V_3 on sufficiently many quadratic observables will eventually resolve all uncertainty about these quantities. Of course, in practice one will have only a finite number of observations, N , with which to update beliefs about the underlying variance structure. The *adjusted expectation* for \mathbf{V} , $E_D(\mathbf{V})$ given the three finite series of quadratic observables, $\{X_t^{(1)}|3 \leq t \leq N\}$, $\{X_t^{(2)}|4 \leq t \leq N\}$ and $\{X_t^{(3)}|5 \leq t \leq N\}$ takes the form

$$E_D(\mathbf{V}) = E(\mathbf{V}) + \text{Cov}(\mathbf{V}, \mathbf{D})[\text{Var}(\mathbf{D})]^\dagger[\mathbf{D} - E(\mathbf{D})] \quad (24)$$

where \mathbf{D} is the $(3N - 9)$ -dimensional vector

$$\mathbf{D} = (X_3^{(1)2}, \dots, X_N^{(1)2}, X_4^{(2)2}, \dots, X_N^{(2)2}, X_5^{(3)2}, \dots, X_N^{(3)2})^T \quad (25)$$

and $[\text{Var}(\mathbf{D})]^\dagger$ denotes the Moore-Penrose generalised inverse of $\text{Var}(\mathbf{D})$. All necessary covariances are given in the previous section and the appendix. Note that there is nothing particularly special about the choice of \mathbf{D} other than the fact that it is one of the smallest choices of \mathbf{D} which will lead to identification of the underlying variance structure. It could be enlarged by introducing terms of the form $X_t^{(4)2}$ etc. This would lead to a richer projection space, and hence more efficient estimates. However, one would have to compute the covariance structure over the extra observables, and this exercise would quickly become computationally unattractive.

7 Example

The theory developed thus far will now be applied to a genuine example arising from research into the forecasting of competitive retail markets. Figure 1 shows a time series for the case sales of the leading brand of cola from a particular wholesale depot in England for the first 200 days of 1995. It is assumed that these sales figures, $\{X_t|1 \leq t \leq 200\}$, follow a second-order locally linear time-series DLM, and so the 3-, 4- and 5-step exchangeable sets of quantities, $\{X_t^{(1)}|3 \leq t \leq 200\}$, $\{X_t^{(2)}|4 \leq t \leq 200\}$ and $\{X_t^{(3)}|5 \leq t \leq 200\}$ are

formed, as described in Section 3.1, and these are shown in Figure 2. According to our model, these sets of quantities are mean zero n -step exchangeable, and so obvious evidence of an underlying mean away from zero, or evidence of long-range dependencies would be evidence against the model. There does not appear to be any obvious discrepancies. Note that short-range dependencies, and dependencies between the series are to be expected.

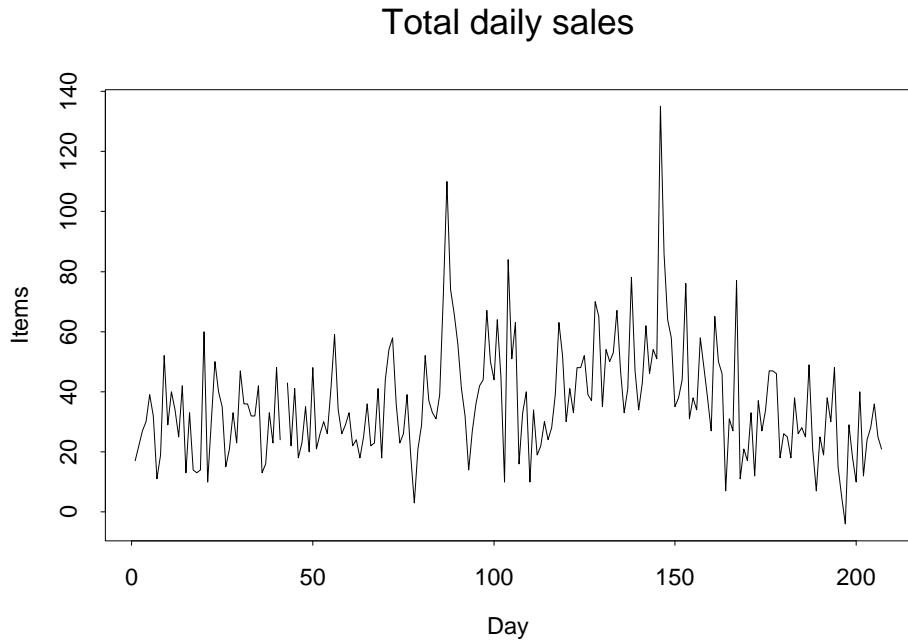


Figure 1: Plot showing the time series of sales for the example

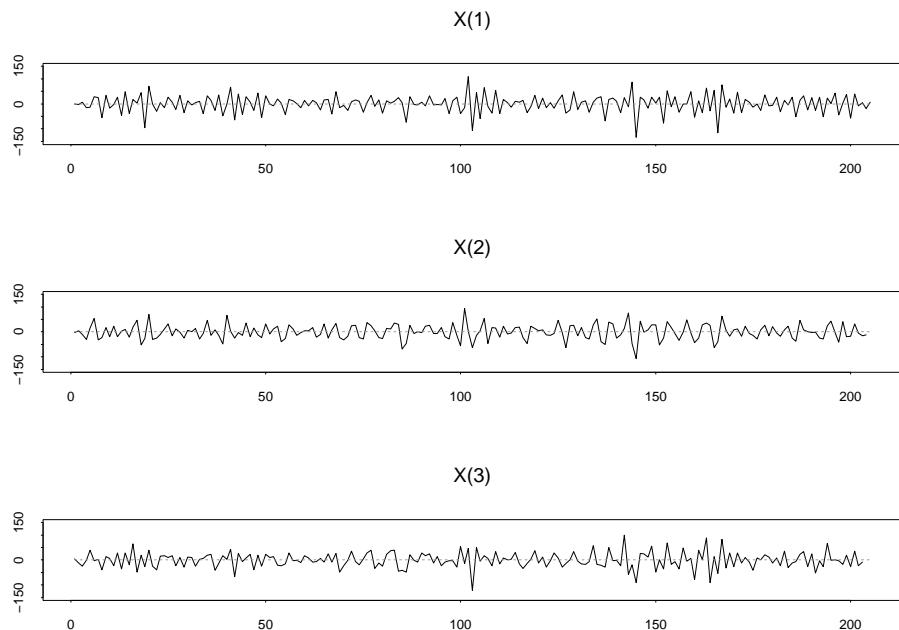


Figure 2: Plot showing the mean zero linear combinations

Next, the quadratic observables $\{X_t^{(1)2} | 3 \leq t \leq 200\}$, $\{X_t^{(2)2} | 4 \leq t \leq 200\}$ and $\{X_t^{(3)2} | 5 \leq t \leq 200\}$ are formed, as described in Section 3.2, and these are shown in Figure 3. Again, due to assumptions about the exchangeability of the quadratic residuals, these series are each n -step exchangeable, and so obvious long-range dependencies would be evidence against our model. Fortunately, none are apparent, although short range dependencies, and dependencies between series are particularly clear in this figure.

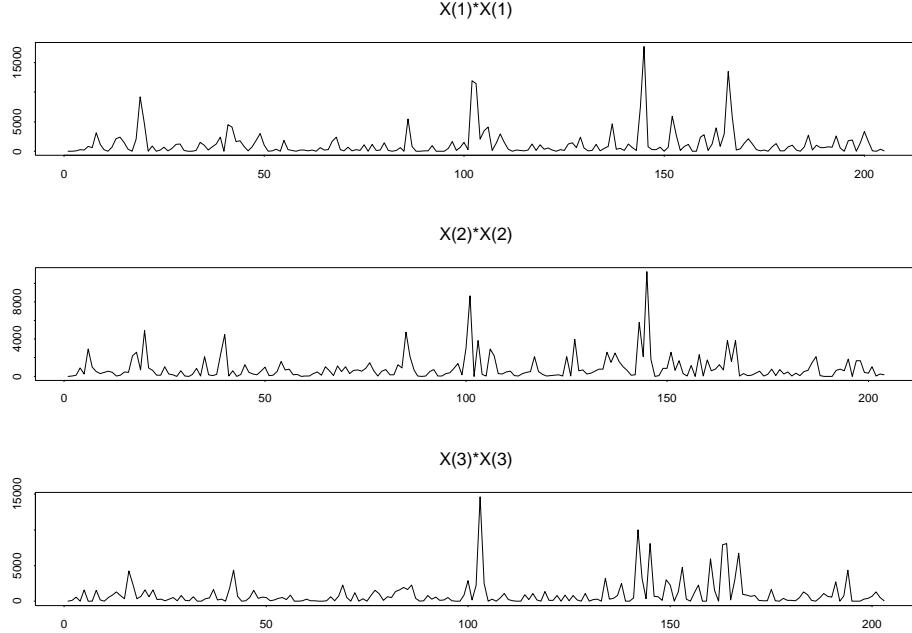


Figure 3: Plot showing quadratic observables

A priori belief specifications are required before analysis can take place. The specifications required for a basic linear analysis of this problem were made as follows.

$$E(M_1) = 20, \text{Var}(M_1) = 20^2, E(N_1) = 0, \text{Var}(N_1) = 3^2 \quad (26)$$

$$E(V_1) = \text{Var}(Y_{1t}) = 5^2, E(V_2) = \text{Var}(Y_{2t}) = 0.2^2, E(V_3) = \text{Var}(Y_{3t}) = 0.1^2, \forall t \quad (27)$$

The above specifications are also precisely those required for a fully Bayesian approach to the analysis of the locally linear time series DLM, together with some distributional assumptions, as described in [8]. The additional specifications required for a quadratic analysis are given below.

$$\text{Var}(V_1) = 5^2, \text{Var}(V_2) = 1^2, \text{Var}(V_3) = 0.2^2 \quad (28)$$

$$\text{Var}(S_{1t}) = 2(5^4), \text{Var}(S_{2t}) = 2(0.2^4), \text{Var}(S_{3t}) = 2(0.1^4), \forall t \quad (29)$$

Note that in this example, for simplicity, the variance specifications for the S_{it} have been made to be consistent with a χ^2 fit for the distribution of Y_{it}^2 , given their underlying mean. Such fitting is discussed in the more general multivariate context in [9]. Note however, that such fitting is not required, and that in any case, assignment of the fourth moments in this way is a much weaker assumption than that of full normality of the Y_{it} .

Clearly, the additional specification burden required in order to carry out variance learning is quite small. Specification of six additional numbers (three, if one is prepared to fit the $\text{Var}(S_{it})$) is all that is required. Note also that since we take a Bayes linear approach, adjustments reduce to the solving of matrix equations, and so the computational requirements are not great.

The Bayes linear computing package, [B/D], was used to analyse the problem. “Elements” corresponding to the linear and quadratic terms were “built”, and expectations and covariances were assigned appropriately, using output from computer algebra systems where necessary. The computer algebra systems were used for numerical substitution of beliefs into the covariance formulae given in the appendix, as well as for the algebraic

derivation of the formulae themselves. Beliefs about V_1 , V_2 and V_3 were then adjusted using the quadratic observables formed from the first 200 observations from the series. The sequence of adjusted expectations, together with corresponding two-standard deviation credibility bounds for the variance components are shown in Figure 4. Notice that beliefs about all of the variances have been revised upwards. In particular, beliefs about the magnitude of V_1 have been revised upwards quite sharply, relative to prior uncertainty. The magnitude of the revisions should be regarded as a diagnostic warning for the high-order variance specification, since the eventual adjusted expectation lies some way outside the *a priori* credibility bounds. The adjustments for the variances were as follows.

$$E_D(V_1) = 171, \quad E_D(V_2) = 4.75, \quad E_D(V_3) = 0.36 \quad (30)$$

We shall see the implications of these revisions for first order adjustments, later. These adjustment sequences may be compared with natural unbiased sample estimates for the variance components, as shown in Figure 5. These estimates can be seen to be unstable, due to the fact that the variance of these estimates is very large, even with 200 observations.

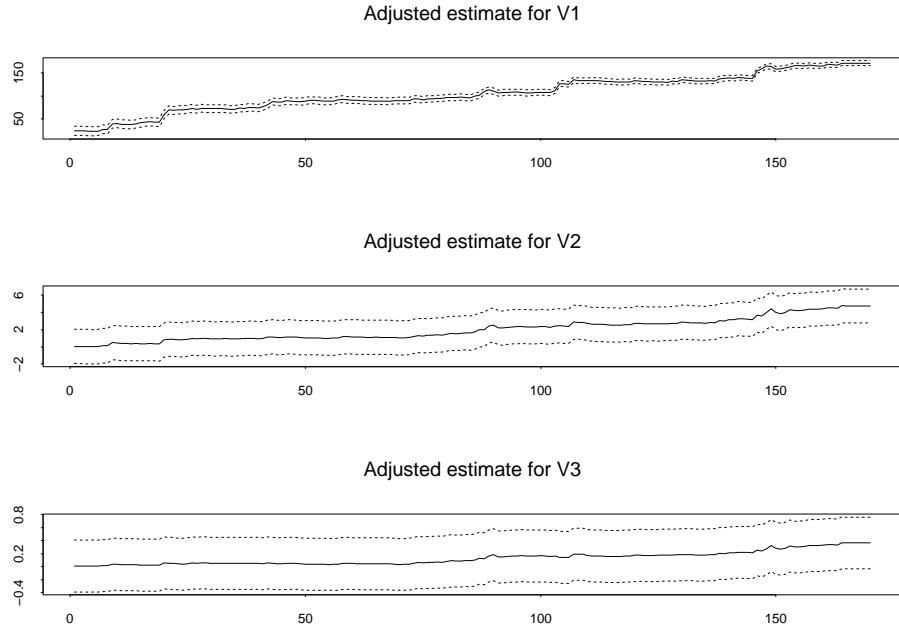


Figure 4: Variance component adjustments

We may now consider the linear forecasting problem, and compare the forecasts of that model using the original, and adjusted variance specifications. Figure 6 shows the first order adjustments, together with a one-step ahead predictive two-standard deviation credibility interval, using the original variance specifications. It appears that the specification for $\text{Var}(Y_{1t})$ may be too small, since considerably more than 5% of observations lie outside the credibility interval. Figure 7 shows the same, using the adjusted variances. The increased specification for $\text{Var}(Y_{1t})$ leads to wider credibility intervals, which can be clearly seen to match better with the forecast performance of the model. Notice also that the increased estimate for V_2 allows the mean parameter, M_t , to adapt more quickly to fluctuations in the data.

8 Conclusions

Appropriate variance specifications are crucial to the performance of dynamic linear models. Even a carefully chosen, and wholly appropriate model will perform very poorly if the variance specifications used are wrong. This fact has long been appreciated, and in [8], methods for the updating of the top-level variance (V_1 in this paper) are discussed. However, such methods remain the state of the art, and yet, not all uncertainty for the top-level variance is resolved, and no methods are given for data-driven learning for other variance

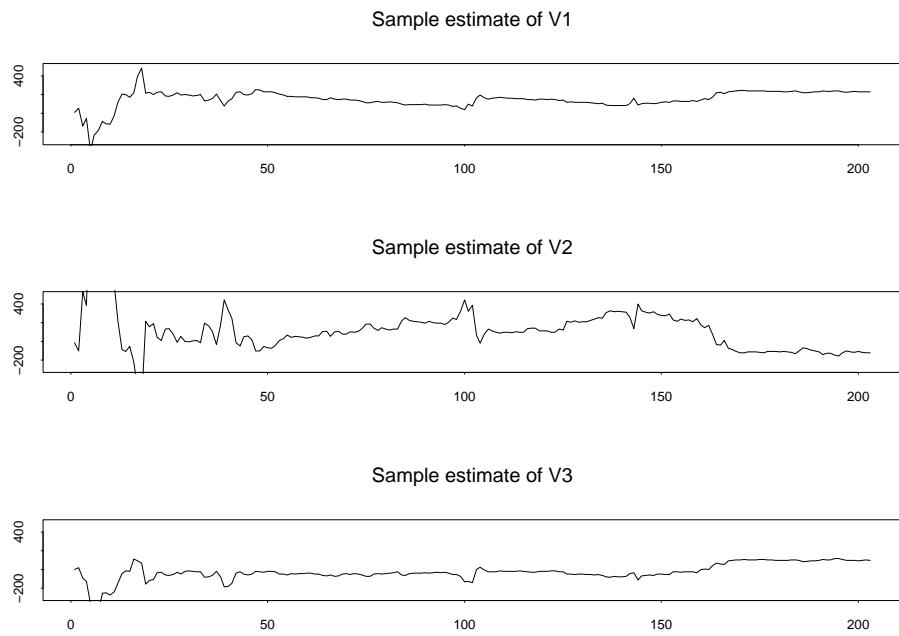


Figure 5: Unbiased sample estimates for the variance components

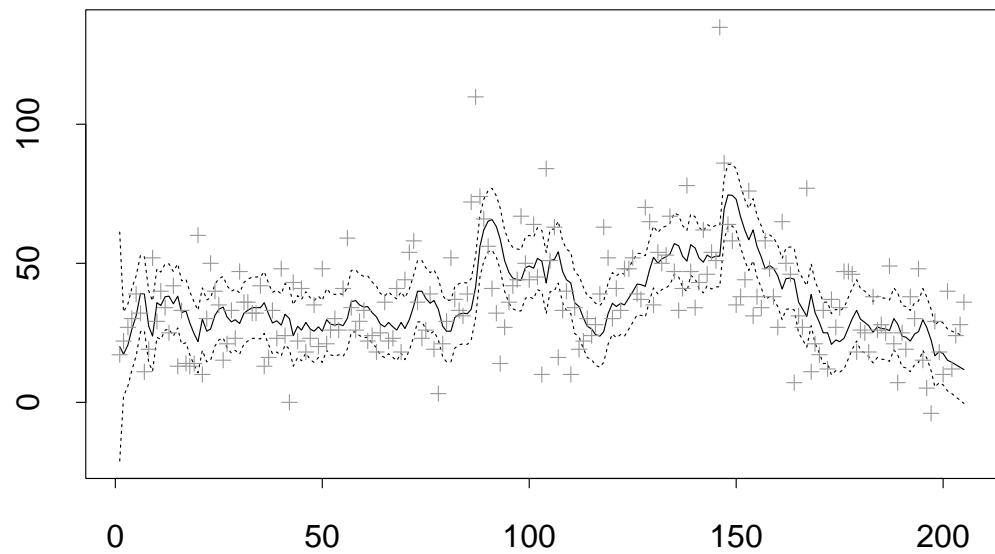


Figure 6: First order adjustments without variance revision

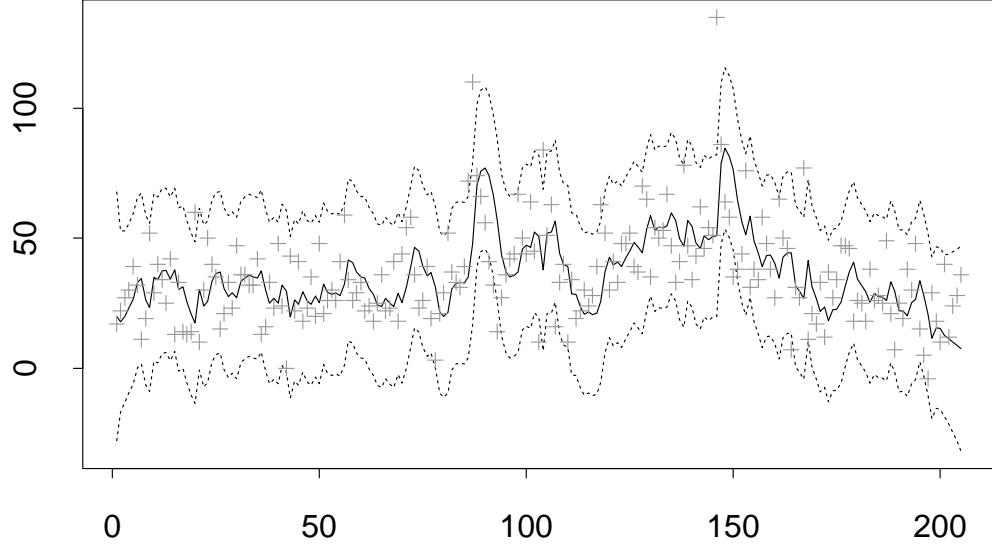


Figure 7: First order adjustments using revised variances

components in the model. Appropriate specifications for other variance components are just as important, and can have just as much effect on the performance of the model (especially for mean-dominated series, such as many financial time series). In the example given in this paper, learning that V_2 was bigger than thought allowed the series parameters to adapt more quickly to the data. Such data-driven learning for the parameter variances is not possible using existing DLM methodology. Further, it is hard to see how the methodology might be adapted to allow such learning. Here, methods are given for learning about all of the variance components for a locally linear time series model, using limited partial prior specifications for the variance components, thus providing a subjective Bayesian, computationally tractable solution to the problem.

9 Acknowledgements

All Bayes linear computations were carried out using the Bayes linear computing package, [B/D], available from [11]. The covariance calculations were carried out using the REDUCE computer algebra system, described in [7], and the MuPAD computer algebra system, described in [2] and [12]. The data for the example was provided by *Positive Concepts Ltd.*

Appendix

The covariances between the quadratic observables themselves are rather complex, and were calculated using computer algebra packages in order to ensure accuracy.

Lemma 3 *The covariance structure over the $X_t^{(1)2}$ terms is given by the following relations.*

$$\begin{aligned} \text{Cov}(X_t^{(1)2}, X_t^{(1)2}) = & \text{Var}(S_{3t}) + 2\text{Var}(S_{2t}) + 18\text{Var}(S_{1t}) + \text{Var}(V_3) + 8\text{Var}(V_2) \\ & + 72\text{Var}(V_1) + 8E(V_3)E(V_2) + 24E(V_3)E(V_1) + 4E(V_2)^2 \end{aligned}$$

$$+48E(V_2)E(V_1) + 36E(V_1)^2, \quad \forall t \geq 3 \quad (31)$$

$$\begin{aligned} \text{Cov}(X_t^{(1)^2}, X_{t-1}^{(1)^2}) &= \text{Var}(S_{2t}) + 8\text{Var}(S_{1t}) + \text{Var}(V_3) + 4\text{Var}(V_2) + 52\text{Var}(V_1) \\ &\quad + 16E(V_2)E(V_1) + 16E(V_1)^2, \quad \forall t \geq 4 \end{aligned} \quad (32)$$

$$\text{Cov}(X_t^{(1)^2}, X_{t-2}^{(1)^2}) = \text{Var}(S_{1t}) + \text{Var}(V_3) + 4\text{Var}(V_2) + 36\text{Var}(V_1), \quad \forall t \geq 5 \quad (33)$$

$$\text{Cov}(X_t^{(1)^2}, X_{t-s}^{(1)^2}) = \text{Var}(V_3) + 4\text{Var}(V_2) + 36\text{Var}(V_1), \quad \forall s \geq 3, \forall t \geq s+3 \quad (34)$$

Lemma 4 The covariance structure over the $X_t^{(2)^2}$ terms is given by the following relations.

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_t^{(2)^2}) &= 2(\text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 2\text{Var}(S_{1t}) + 4\text{Var}(V_3) + 4\text{Var}(V_2) \\ &\quad + 20\text{Var}(V_1) + 2E(V_3)^2 + 8E(V_3)E(V_2) + 16E(V_3)E(V_1) + 2E(V_2)^2 \\ &\quad + 16E(V_2)E(V_1) + 12E(V_1)^2), \quad \forall t \geq 4 \end{aligned} \quad (35)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t-1}^{(2)^2}) &= \text{Var}(S_{3t}) + 3\text{Var}(S_{1t}) + 4\text{Var}(V_3) + 4\text{Var}(V_2) + 12\text{Var}(V_1) \\ &\quad - 4E(V_3)E(V_1) - 4E(V_1)^2, \quad \forall t \geq 5 \end{aligned} \quad (36)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t-2}^{(2)^2}) &= \text{Var}(S_{2t}) + 2\text{Var}(S_{1t}) + 4\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 8E(V_2)E(V_1) + 4E(V_1)^2, \quad \forall t \geq 6 \end{aligned} \quad (37)$$

$$\text{Cov}(X_t^{(2)^2}, X_{t-3}^{(2)^2}) = \text{Var}(S_{1t}) + 4\text{Var}(V_3) + 4\text{Var}(V_2) + 16\text{Var}(V_1), \quad \forall t \geq 7 \quad (38)$$

$$\text{Cov}(X_t^{(2)^2}, X_{t-s}^{(2)^2}) = 4(\text{Var}(V_3) + \text{Var}(V_2) + 4\text{Var}(V_1)), \quad \forall s \geq 4, \forall t \geq s+4 \quad (39)$$

Lemma 5 The covariance structure over the $X_t^{(3)^2}$ terms is given by the following relations.

$$\begin{aligned} \text{Cov}(X_t^{(3)^2}, X_t^{(3)^2}) &= 3\text{Var}(S_{3t}) + 2\text{Var}(S_{2t}) + 4\text{Var}(S_{1t}) + 21\text{Var}(V_3) + 8\text{Var}(V_2) \\ &\quad + 40\text{Var}(V_1) + 12E(V_3)^2 + 24E(V_3)E(V_2) + 48E(V_3)E(V_1) + 4E(V_2)^2 \\ &\quad + 32E(V_2)E(V_1) + 24E(V_1)^2, \quad \forall t \geq 5 \end{aligned} \quad (40)$$

$$\begin{aligned} \text{Cov}(X_t^{(3)^2}, X_{t-1}^{(3)^2}) &= 2\text{Var}(S_{3t}) + 2\text{Var}(S_{1t}) + 13\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 4E(V_3)^2 - 16E(V_3)E(V_1) + 4E(V_1)^2, \quad \forall t \geq 6 \end{aligned} \quad (41)$$

$$\begin{aligned} \text{Cov}(X_t^{(3)^2}, X_{t-2}^{(3)^2}) &= \text{Var}(S_{3t}) + \text{Var}(S_{1t}) + 9\text{Var}(V_3) + 4\text{Var}(V_2) + 16\text{Var}(V_1) \\ &\quad + 4E(V_3)E(V_1), \quad \forall t \geq 7 \end{aligned} \quad (42)$$

$$\begin{aligned} \text{Cov}(X_t^{(3)^2}, X_{t-3}^{(3)^2}) &= \text{Var}(S_{2t}) + 2\text{Var}(S_{1t}) + 9\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 8E(V_2)E(V_1) + 4E(V_1)^2, \quad \forall t \geq 8 \end{aligned} \quad (43)$$

$$\text{Cov}(X_t^{(3)^2}, X_{t-4}^{(3)^2}) = \text{Var}(S_{1t}) + 9\text{Var}(V_3) + 4\text{Var}(V_2) + 16\text{Var}(V_1), \quad \forall t \geq 9 \quad (44)$$

$$\text{Cov}(X_t^{(3)^2}, X_{t-s}^{(3)^2}) = 9\text{Var}(V_3) + 4\text{Var}(V_2) + 16\text{Var}(V_1), \quad \forall s \geq 5, \forall t \geq s+5 \quad (45)$$

Lemma 6 The covariance structure between the $X_t^{(1)^2}$ and $X_t^{(2)^2}$ terms is given by the following relations.

$$\text{Cov}(X_t^{(1)^2}, X_{t+s}^{(2)^2}) = 2(\text{Var}(V_3) + 2\text{Var}(V_2) + 12\text{Var}(V_1)), \quad \forall t \geq 3, \forall s \geq 4 \quad (46)$$

$$\text{Cov}(X_t^{(1)^2}, X_{t+3}^{(2)^2}) = \text{Var}(S_{1t}) + 2\text{Var}(V_3) + 4\text{Var}(V_2) + 24\text{Var}(V_1), \quad \forall t \geq 3 \quad (47)$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t+2}^{(2)^2}) &= \text{Var}(S_{2t}) + 5\text{Var}(S_{1t}) + 2\text{Var}(V_3) + 4\text{Var}(V_2) + 32\text{Var}(V_1) \\ &\quad + 12E(V_2)E(V_1) + 8E(V_1)^2, \quad \forall t \geq 3\end{aligned}\tag{48}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t+1}^{(2)^2}) &= \text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 6\text{Var}(S_{1t}) + 2\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 4E(V_3)E(V_2) + 8E(V_3)E(V_1) + 8E(V_2)E(V_1) - 4E(V_1)^2, \quad \forall t \geq 3\end{aligned}\tag{49}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_t^{(2)^2}) &= \text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 6\text{Var}(S_{1t}) + 2\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 4E(V_3)E(V_2) + 8E(V_3)E(V_1) + 8E(V_2)E(V_1) - 4E(V_1)^2, \quad \forall t \geq 4\end{aligned}\tag{50}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t-1}^{(2)^2}) &= \text{Var}(S_{2t}) + 5\text{Var}(S_{1t}) + 2\text{Var}(V_3) + 4\text{Var}(V_2) + 32\text{Var}(V_1) \\ &\quad + 12E(V_2)E(V_1) + 8E(V_1)^2, \quad \forall t \geq 5\end{aligned}\tag{51}$$

$$\text{Cov}(X_t^{(1)^2}, X_{t-2}^{(2)^2}) = \text{Var}(S_{1t}) + 2\text{Var}(V_3) + 4\text{Var}(V_2) + 24\text{Var}(V_1), \quad \forall t \geq 6\tag{52}$$

$$\text{Cov}(X_t^{(1)^2}, X_{t-s}^{(2)^2}) = 2(\text{Var}(V_3) + 2\text{Var}(V_2) + 12\text{Var}(V_1)), \quad \forall s \geq 3, \forall t \geq s+4\tag{53}$$

Lemma 7 The covariance structure between the $X_t^{(1)^2}$ and $X_t^{(3)^2}$ terms is given by the following relations.

$$\text{Cov}(X_t^{(1)^2}, X_{t+s}^{(3)^2}) = 3\text{Var}(V_3) + 4\text{Var}(V_2) + 24\text{Var}(V_1), \quad \forall s \geq 5, \forall t \geq 3\tag{54}$$

$$\text{Cov}(X_t^{(1)^2}, X_{t+4}^{(3)^2}) = \text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) + 24\text{Var}(V_1), \quad \forall t \geq 3\tag{55}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t+3}^{(3)^2}) &= \text{Var}(S_{2t}) + 5\text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) + 32\text{Var}(V_1) \\ &\quad + 12E(V_2)E(V_1) + 8E(V_1)^2, \quad \forall t \geq 3\end{aligned}\tag{56}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t+2}^{(3)^2}) &= \text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 5\text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) \\ &\quad + 32\text{Var}(V_1) + 4E(V_3)E(V_2) + 12E(V_3)E(V_1) \\ &\quad + 12E(V_2)E(V_1) + 8E(V_1)^2, \quad \forall t \geq 3\end{aligned}\tag{57}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t+1}^{(3)^2}) &= \text{Var}(S_{3t}) + 2\text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) + 28\text{Var}(V_1) \\ &\quad - 8E(V_3)E(V_1) + 4E(V_1)^2, \quad \forall t \geq 4\end{aligned}\tag{58}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_t^{(3)^2}) &= \text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 5\text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) \\ &\quad + 32\text{Var}(V_1) + 4E(V_3)E(V_2) + 12E(V_3)E(V_1) \\ &\quad + 12E(V_2)E(V_1) + 8E(V_1)^2, \quad \forall t \geq 5\end{aligned}\tag{59}$$

$$\begin{aligned}\text{Cov}(X_t^{(1)^2}, X_{t-1}^{(3)^2}) &= \text{Var}(S_{2t}) + 5\text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) + 32\text{Var}(V_1) \\ &\quad + 12E(V_2)E(V_1) + 8E(V_1)^2, \quad \forall t \geq 6\end{aligned}\tag{60}$$

$$\text{Cov}(X_t^{(1)^2}, X_{t-2}^{(3)^2}) = \text{Var}(S_{1t}) + 3\text{Var}(V_3) + 4\text{Var}(V_2) + 24\text{Var}(V_1), \quad \forall t \geq 7\tag{61}$$

$$\text{Cov}(X_t^{(1)^2}, X_{t-s}^{(3)^2}) = 3\text{Var}(V_3) + 4\text{Var}(V_2) + 24\text{Var}(V_1), \quad \forall s \geq 3, \forall t \geq s+4\tag{62}$$

Lemma 8 The covariance structure between the $X_t^{(2)^2}$ and $X_t^{(3)^2}$ terms is given by the following relations.

$$\text{Cov}(X_t^{(2)^2}, X_{t+s}^{(3)^2}) = 2(3\text{Var}(V_3) + 2\text{Var}(V_2) + 8\text{Var}(V_1)), \quad \forall t \geq 4, \forall s \geq 5\tag{63}$$

$$\text{Cov}(X_t^{(2)^2}, X_{t+4}^{(3)^2}) = \text{Var}(S_{1t}) + 6\text{Var}(V_3) + 4\text{Var}(V_2) + 16\text{Var}(V_1), \quad \forall t \geq 4\tag{64}$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t+3}^{(3)^2}) &= \text{Var}(S_{2t}) + 2\text{Var}(S_{1t}) + 6\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 8E(V_2)E(V_1) + 4E(V_1)^2, \quad \forall t \geq 4 \end{aligned} \quad (65)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t+2}^{(3)^2}) &= \text{Var}(S_{3t}) + 2\text{Var}(S_{1t}) + 6\text{Var}(V_3) + 4\text{Var}(V_2) + 12\text{Var}(V_1) \\ &\quad - 4E(V_1)^2, \quad \forall t \geq 4 \end{aligned} \quad (66)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t+1}^{(3)^2}) &= 2\text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 3\text{Var}(S_{1t}) + 10\text{Var}(V_3) + 4\text{Var}(V_2) \\ &\quad + 12\text{Var}(V_1) + 4E(V_3)^2 + 8E(V_3)E(V_2) + 8E(V_3)E(V_1) \\ &\quad + 4E(V_2)E(V_1) - 4E(V_1)^2, \quad \forall t \geq 4 \end{aligned} \quad (67)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_t^{(3)^2}) &= 2\text{Var}(S_{3t}) + \text{Var}(S_{2t}) + 3\text{Var}(S_{1t}) + 10\text{Var}(V_3) + 4\text{Var}(V_2) \\ &\quad + 12\text{Var}(V_1) + 4E(V_3)^2 + 8E(V_3)E(V_2) + 8E(V_3)E(V_1) \\ &\quad + 4E(V_2)E(V_1) - 4E(V_1)^2, \quad \forall t \geq 5 \end{aligned} \quad (68)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t-1}^{(3)^2}) &= \text{Var}(S_{3t}) + 2\text{Var}(S_{1t}) + 6\text{Var}(V_3) + 4\text{Var}(V_2) + 12\text{Var}(V_1) \\ &\quad - 4E(V_1)^2, \quad \forall t \geq 6 \end{aligned} \quad (69)$$

$$\begin{aligned} \text{Cov}(X_t^{(2)^2}, X_{t-2}^{(3)^2}) &= \text{Var}(S_{2t}) + 2\text{Var}(S_{1t}) + 6\text{Var}(V_3) + 4\text{Var}(V_2) + 20\text{Var}(V_1) \\ &\quad + 8E(V_2)E(V_1) + 4E(V_1)^2, \quad \forall t \geq 7 \end{aligned} \quad (70)$$

$$\text{Cov}(X_t^{(2)^2}, X_{t-3}^{(3)^2}) = \text{Var}(S_{1t}) + 6\text{Var}(V_3) + 4\text{Var}(V_2) + 16\text{Var}(V_1), \quad \forall t \geq 8 \quad (71)$$

$$\text{Cov}(X_t^{(2)^2}, X_{t-s}^{(3)^2}) = 2(3\text{Var}(V_3) + 2\text{Var}(V_2) + 8\text{Var}(V_1)), \quad \forall s \geq 4, \forall t \geq s+5 \quad (72)$$

Proof

All results were first derived using a simple program for the REDUCE computer algebra system. The output was used for further computations, and also passed through a filter which type-set the results accurately for L^AT_EX. All calculations and computations were verified by re-coding the entire problem in a very different way for the MuPAD computer algebra system. The author would be willing to e-mail the REDUCE and/or MuPAD programs to any interested parties. \square

Note that the “far-away” covariances, (34), (39), (45), (46), (53), (54), (62), (63) and (72) may be deduced directly using Lemma 1 and the n -step exchangeability representation theorem. Unfortunately, the other covariances are not so amenable to such an approach.

References

- [1] M. Farrow and M. Goldstein. Bayes linear methods for grouped multivariate repeated measurement studies with application to crossover trials. *Biometrika*, 80(1):39–59, 1993.
- [2] B. Fuchssteiner et al. *MuPAD Tutorial*. Birkhäuser, Basel, 1994.
- [3] M. Goldstein. Exchangeable belief structures. *J. Amer. Statist. Ass.*, 81:971–976, 1986.
- [4] M. Goldstein. Revising exchangeable beliefs: subjectivist foundations for the inductive argument. In P. Freeman and A.F.M. Smith, editors, *Aspects of Uncertainty: A Tribute to D. V. Lindley*. Wiley, 1994.
- [5] M. Goldstein. Prior inferences for posterior judgements. In M. L. D. Chiara et al., editors, *10th International Congress of Logic, Methodology and Philosophy of Science*. Kluwer, to appear, 1996.
- [6] M. Goldstein and D. A. Wooff. Bayes linear computation: concepts, implementation and programming environment. *Statistics and Computing*, 5:327–341, 1995.

- [7] G. Rayna. *REDUCE: Software for algebraic computation*. Springer-Verlag, 1987.
- [8] M. West and P. J. Harrison. *Bayesian forecasting and dynamic models*. Springer, New York, 1989.
- [9] D. J. Wilkinson. *Bayes linear covariance matrix adjustment*. PhD thesis, University of Durham, 1995.
- [10] D. J. Wilkinson and M. Goldstein. Bayes linear covariance matrix adjustment for multivariate dynamic linear models. *Bayesian analysis e-print*, URL: <http://xxx.lanl.gov/abs/bayes-an/9506002>, 1995.
- [11] D. A. Wooff and M. Goldstein. [B/D] — Beliefs adjusted by Data: Bayes linear programming language. *Internet site* <http://fourier.dur.ac.uk:8000/stats/bd/>, 1995.
- [12] Paul Zimmermann. Wester's test suite in MuPAD 1.2.2. *Computer Algebra Nederland Nieuwsbrief*, (14):53–64, April 1995.